Statistical Decision Making for Authentication and Intrusion Detection

Christos Dimitrakakis
Informatics Institute,
University of Amsterdam,
Science Park 107, NL-1098XG Amsterdam
c.dimitrakakis@uva.nl

Aikaterini Mitrokotsa
Faculty of EEMCS
Delft University of Technology
Mekelweg 4, NL-2628CD Delft
A.Mitrokotsa@TUDelft.nl

October 6, 2009

Abstract

User authentication and intrusion detection differ from standard classification problems in that while we have data generated from legitimate users, impostor or intrusion data is scarce or non-existent. We review existing techniques for dealing with this problem and propose a novel alternative based on a principled statistical decision-making view point. We examine the technique on a toy problem and validate it on complex real-world data from an RFID based access control system. The results indicate that it can significantly outperform the classical world model approach. The method could be more generally useful in other decision-making scenarios where there is a lack of adversary data.

1 Introduction

Classification is the problem of categorising data in one of two or more possible classes. In the classical *supervised learning* framework, examples of each class have already been obtained and the task of the decision maker is to accurately categorise new observations, whose class is unknown. The accuracy is either measured in terms of the rate of misclassification, or in terms of the average cost, for problems where different types of errors carry different costs. In that setting, the problem has three phases: (a) the collection of training data, (b) the estimation of a decision rule based on the training data and (c) the application

of the decision rule to new data. Typically, the decision rule remains fixed after the second step. Thus, the problem becomes that of finding the decision rule with minimum risk from the training data.

Unfortunately, some problems are structured in such a way that it is not possible to obtain data from all categories to form the decision rule. Novelty detection, user authentication, network intrusion detection and spam filtering all belong to this type of decision problems: while the data of the "normal" class is relatively easily characterised, the data of the other class which we wish to detect is not. This is partially due to the potentially adversarial nature of the process that generates the data of the alternative class. As an example, consider being asked to decide whether a particular voice sample belongs to a specific person, given a set of examples of his voice and your overall experience concerning the voices of other persons.

In this paper, we shall employ two conceptual classes: the "user" and the "adversary". The main distinction is that while we shall always have examples of instances of the user class, we may not have any data from the adversary class.

This problem is alleviated in authentication settings, where we must separate accesses by a specific user from accesses by an adversary. Such problems contain additional information: data which we have obtained from other people. This can be used to create a *world model*, which can then act as an adversary model, and has been used with state-of-the-art results in authentication [5, 12, 20].

Since there is no explicit adversary model, the probability of an attack cannot be estimated. Our main contribution is a decision making principle which employs a *pessimistic estimate* on the probability of an attack. Intuitively, this is done by conditioning the adversary model on the *current* observations, whose class is unknown. This enables us to place an upper bound on the probability of the adversary class, in a *Bayesian* framework. To the best of our knowledge, this is the first time that such a Bayesian worst-case approach has been described in the literature. The proposed method is compared with both an *oracle* and the *world model* approach on a test-bench. This shows that our approach can outperform the world model under a variety of conditions. This result is validated on the real-world problem of detecting unauthorised accesses in a building.

The remainder of this section discusses related work. The model framework is introduced in Sec. 2, with the proposed Bayesian estimates discussed in Sec. 2.2 and methods for estimating the prior in Sec. 2.3. The conclusion is preceded by Sec. 3, which presents experiments and results.

1.1 Related work

Classification algorithms have been extensively used for the detection of intrusions in wired [4, 17] and wireless [8, 14] networks. Their main disadvantage is that labelled normal and attack data must be available for training. After the training phase, the classifier's learnt model will be used to predict the labels of new unknown data. However, such data is very hard to obtain and often unreliable. Finally, there will always exist new unknown attacks for which training

data are not available at all.

Outlier detection [3, 19] and clustering [18] use unlabelled data and are in principle able to detect unknown types of attacks. The main disadvantage is that no explicit adversarial model is employed.

An alternative framework is the world model approach [5, 12, 20]. This is extensively used in speech and image authentication problems, where data from a considerable number of users are collected to create a world model (also called a universal background model). This approach is closely related to the model examined in this paper, since it originates in the seminal work of [13], who employed an empirical Bayes technique for estimating a prior over models. Thus, the world model is a distribution over models, although due to computational considerations a point estimate is used instead in practice [20].

The adversary may actively try to avoid detection, through knowledge of the detection method. In essence, this changes the setting from a statistical to an adversarial one. For such problems, game theoretic approaches are frequently used. Dalvi et al. [6] investigated the adversarial classification problem as a two-person game. More precisely, they examined the optimal strategy of an adversary against a standard (adversary-unaware) classifier as well as that of a classifier (adversary-aware) against a rational adversary. This was under the assumption that the adversary has complete knowledge of the detection algorithm. In a similar vein, Lowd et al. [15] have investigated algorithms for reverse engineering linear classifiers. This allows them to retrieve sufficient information to mount effective attacks.

In our paper we do not consider repeated interactions and thus we do not follow a game-theoretic approach. We instead consider how to model the adversary, when we have a lot of data from legitimate users, but no data from the adversary. Our main contribution is a Bayesian method for calculating a subjective upper bound on attack probabilities without any knowledge of the adversary model. This can be obtained simply by using the current (unlabelled) observations to create a worst-case (or more generally pessimistic) model of the adversary.[1] This is done by conditioning the prior over adversary models according to new (unlabelled) observations.

However, in order to control overfitting, we first condition the adversary model's prior on the data of the remaining population of users. This results in an empirical Bayes estimate of the prior [21], which is what the world model approach essentially is [13]. The prior then acts as a soft constraint when selecting the worst-case adversary model.

It is worthwhile to note that the problem of constructing a model for a class with no data is related to the problem of null hypothesis testing, for which similar ideas have appeared. For example, [10] explored the idea of constructing a maximum likelihood estimate from the obsrvations and using this as the alternative hypothesis. More sophisticated examples for simple parametric problems were examined in [2]. This involved selecting the worst-case prior from a given class of priors in order to be maximally pessimistic about the null hypothesis.

[1] Some simpler alternative approaches are explored in an accompanying technical report [9].

Our approach is similar in spirit, but the application and technical details are substantially different.

Our final contribution is an experimental analysis on a synthetic problem, as well as on some real-world data, with promising results: we show that the widely used *world model* approach cannot outperform the proposed model.

2 The proposed model framework

In the framework we consider, we assume that the set of all possible models is $\mathcal{M}$. Each model μ in $\mathcal{M}$ is associated with a probability measure over the set of observations $\mathcal{X}$, which will be denoted by $\mu(x)$ for $x \in \mathcal{X}$, $\mu \in \mathcal{M}$, so long as there is no ambiguity. We must decide whether some observations $x \in \mathcal{X}$, have been generated by a model q (the user) or a model w (the adversary) in $\mathcal{M}$. Throughout the paper, we assume a prior probability of the user having generated the data, $\mathbf{P}(q)$, with a complementary prior $\mathbf{P}(w) = 1 - \mathbf{P}(q)$, for the adversary.

In the easiest scenario, we have perfect knowledge of $q, w \in \mathcal{M}$. It is then trivial to calculate the probability $\mathbf{P}(q|x)$ that the user q has generated the data x. This is the *oracle* decision rule, defined in section 2.1. This is not a realisable rule, as although we could accurately estimate q with enough data, in general there is no way to estimate the adversary model w.

We thus consider the case where the user model is known and where we are given a prior density $\xi(w)$ over the possible adversary models $w \in \mathcal{M}$. Currently seen observations are then used to form a pessimistic posterior ξ' for the adversary. This is explained in Section 2.2.

Section 2.3 discusses the more practical case where neither the user model q, nor a prior ξ over models $\mathcal{M}$ are known, but must be estimated from data. More precisely, the section discusses methods for utilising other user data to obtain a prior distribution over models. This amounts to an empirical Bayes estimate of the prior distribution [21]. It is then possible to estimate q by conditioning the prior on the user data. This is closely related to the adapted world model approach [20], used in authentication applications, which however, usually employs a point approximation to the prior [5].

2.1 The oracle decision rule

We shall measure the performance of all the models against that of the *oracle* decision rule. The oracle enjoys perfect information about the distribution of both the user and the adversary, and thus knows both q and w, as well as the *a priori* probability of an attack, $\mathbf{P}(w)$. On average, no other decision rule can do better.

More precisely, let $\mathcal{M}$ be the space of all models. Let the adversary's model be w and the user's model be q, with $q, w \in \mathcal{M}$. Given some data x, we would like to determine the probability that the data x has been generated by the user, $\mathbf{P}(q|x)$. The oracle model has knowledge of w, q and $\mathbf{P}(q)$, so using Bayes'

rule we obtain:

$$\mathbf{P}(q|x) = \frac{q(x)\mathbf{P}(q)}{q(x)\mathbf{P}(q) + w(x)(1 - \mathbf{P}(q))}. \tag{1}$$

However, we usually have uncertainty about both the adversary and the user model. Concerning the adversary, the uncertainty is much more pronounced. The next section examines a model for the probability of an attack when the user model is perfectly known but we only have a prior $\xi(w)$ for the adversary model.

2.2 Bayesian adversary model

We can use a subjective prior probability $\xi(w)$ over possible adversary models, to calculate the probability of observations given that they have been generated by the adversary: $\xi(x) = \int_{\mathcal{M}} w(x)\xi(w)dw$.[2] Given a user model q, we can express the probability of the user q given the observations x under the belief ξ as:

$$\xi(q|x) \triangleq \mathbf{P}(q|x, \xi) = \frac{q(x)\mathbf{P}(q)}{q(x)\mathbf{P}(q) + \xi(x)\,(1 - \mathbf{P}(q))}. \tag{2}$$

The difference with (1) is that, instead of $w(x)$, we use the marginal density $\xi(x)$. If $\xi(w)$ represents our subjective belief about the adversary model w, then (2) can be seen as the Bayesian equivalent of the world model approach, where the prior over w plays the role of the world model. Now let: $\xi'(w) \triangleq \xi(w|x)$ be the model posterior for some observations x. We shall need the following lemma:

Lemma 2.1. *For any probability measure ξ on $\mathcal{M}$, where $\mathcal{M}$ is a space of probability distributions on $\mathcal{X}$, such that each $\mu \in \mathcal{M}$ defines a probability (density) $\mu(x)$ with $x \in \mathcal{X}$, with admissible posteriors $\xi'(\mu) \triangleq \xi(\mu|x)$, the marginal likelihood satisfies: $\xi'(x) \geq \xi(x)$, $\forall x \in \mathcal{X}$.*

A simple proof, using the Cauchy-Schwarz inequality on the norm induced by the measure ξ, is presented in the Appendix. From the above lemma, it immediately follows that:

$$\xi(q|x) \geq \xi'(q|x) = \frac{q(x)\mathbf{P}(q)}{q(x)\mathbf{P}(q) + (1 - \mathbf{P}(q)) \int_{\mathcal{M}} w(x)\xi'(w)\,dw}, \tag{3}$$

since $\xi'(x) = \int w(x)\xi'(w)\,dw \geq \int w(x)\xi'(w)\,dw = \xi(x)$. Thus (3) gives us a subjective upper bound on the probability of the data x having been generated by the adversary. This bound can then be used to make decisions. Finally, note that we can form $\xi'(w)$ on a *subset* of x. This possibility is explored in the experiments.

[2] Here we used the fact that $\xi(x|w) = w(x)$, since the probability of the observations given a specific model w no longer depends on our belief ξ about which model w is correct.

2.3 Prior and user model estimation

Specifically for user authentication, we have data from two sources. The first is data collected from the user which we wish to identify. The second is data collected from other persons.[3] The i-th person can be fully specified in terms of a model $\mu_i \in \mathcal{M}$, with μ_i drawn from some unknown distribution γ over $\mathcal{M}$. If we had the models $\mu_i \in \mathcal{M}$ for all the other people in our dataset, then we could obtain an *empirical* estimate $\hat{\gamma}$ of the prior distribution of models. Empirical Bayes methods for prior estimation [21] extend this procedure to the case where we only observe $x \sim \mu_i$, data drawn from the model μ_i.

Let us now apply this prior over models to the estimation of the posterior over models for some user. Given an estimate $\hat{\gamma}$ of γ, and some data $x \sim \mu$ from the user, and assuming that $\mu \sim \gamma$, we can form a posterior for μ using Bayes rule: $\hat{\gamma}(\mu|x) = \mu(x)\xi(\mu)/\int_{\mathcal{M}} \hat{\gamma}(x|\mu)\xi(d\mu)$, over all $\mu \in \mathcal{M}$. For a specific user k with data x_k, we write the posterior as $\psi_k(\mu) \triangleq \hat{\gamma}(\mu|x_k)$. Whenever we must decide the class of a new observation x, we set the prior over the adversary models to $\xi = \hat{\gamma}$ and then condition on part, or all, of x to obtain the posterior $\xi'(w)$. We then calculate

$$\mathbf{P}(q_k|x, \xi', \psi_k) = \frac{\psi_k(x)\mathbf{P}(q_k)}{[\psi_k(x)\mathbf{P}(q_k) + (1 - \mathbf{P}(q_k))\xi'(x)]}, \tag{4}$$

the posterior probability of the k-th user given the observations x and our beliefs ξ' and ψ_k over adversary and user *models* respectively. When $\xi' = \xi$, we obtain an equivalent to the world model approach of [20], which is an approximate form of the empirical Bayes procedure suggested in [13].

2.3.1 Prior estimation for multinomial models

For discrete observations, we can consider multinomial distributions drawn from a Dirichlet density, and use a maximum likelihood estimate based on Polya distributions for γ. More specifically, we use the fixed point approach suggested in [16] to estimate Dirichlet parameters Φ from a set of multinomial observations.

To make this more concrete, consider multinomial observations of degree K. Our initial belief $\xi(\mu)$ is a Dirichlet prior with parameters $\Phi \triangleq (\phi_1, \ldots, \phi_K)$ over models: $\xi(\mu) = \frac{1}{B(\Phi)} \prod_{i=1}^{K} \mu_i^{\phi_i - 1}$, which is conjugate to the multinomial [7]. Given a sequence of observations $x_1, \ldots, x_n$, with $x_t \in 1, \ldots K$, where each outcome i has fixed probability μ_i, then $c_i = \sum_{t=1}^{n} \mathbb{I}(x_t = i)$, where $\mathbb{I}$ is an indicator function, is multinomial and the posterior distribution over the parameters μ_i is also Dirichlet with parameters $\phi_i' = \phi_i + c_i$. The approach suggested in [16] uses the following fixed point iteration for the parameters: $\phi_i^{new} = \phi_i \frac{\sum_k \Psi(c_{ik} + \phi_i) - \Psi(\phi_i)}{\sum_k \Psi(c_k + \sum_i \phi_i) - \Psi(\sum_i a_i)}$, where $\Psi(\cdot)$ is the digamma function.

[3] These are not necessarily other users.

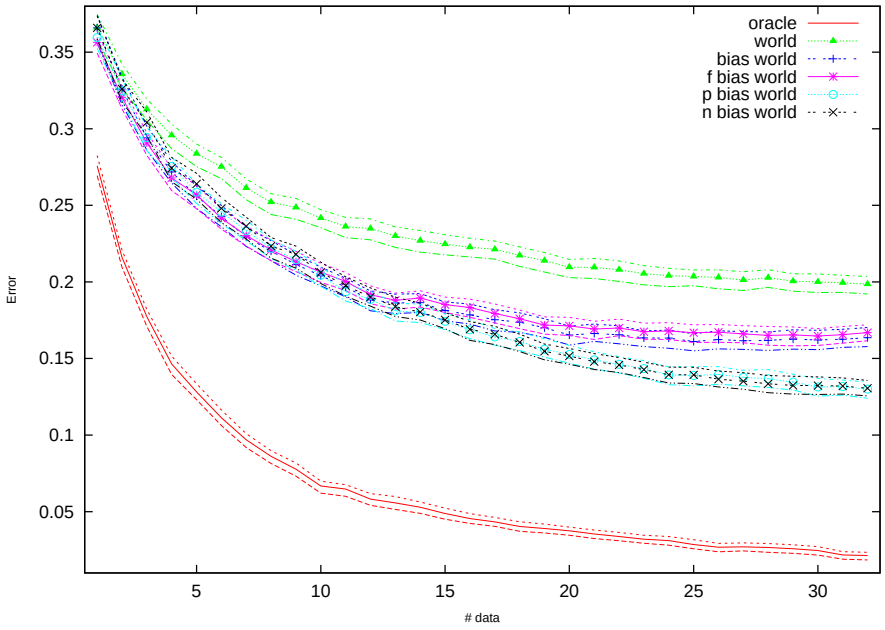

Figure 1: The evolution of error rates as more data becomes available, when the user model and prior are estimated. The points indicate means from 10^4 runs and the lines top and bottom 5% percentiles from a bootstrap sample.

3 Experimental evaluation

We have performed a number of experiments in order to evaluate the proposed approach and compared it to the full Bayesian version of the well-known world model approach. We performed a set of experiments on synthetic data, and another set of experiments on real data.

For the synthetic experiments, we assume multinomial models, but rather than knowing γ, we use data from other users to form an empirical estimate $\hat{\gamma}$, as described in Sec. 2.3.1. Furthermore, q is itself unknown and is estimated via Bayesian updating from $\hat{\gamma}$ and some data specific to the user. We then compare the oracle and the world model approach (based on γ) with a number of differently biased adversary models. The world model is based on the estimate $\hat{\gamma}$. The adversary model uses the world model $\hat{\gamma}$ as the adversary prior (ξ).

The second group concerns experiments on data gathered from an access control system. The data has been discretized into 1320 integer variables, in order for it to be modelled with multinomials. The models are of course not available so we must estimate the priors: The data of a subset of users is used to estimate $\hat{\gamma}$. The remaining users alternatively take on the roles of legitimate users and adversaries.

We compare the following types of models, which correspond to the legends

in the figures of the experimental results. (a) The **oracle** model, which enjoys perfect information concerning adversary and user distributions. (b) The **world** model, which uses the prior over user models as a surrogate for the adversary model. (c) The **bias world** model, which uses all but the last observation to obtain a posterior over adversary models, and similarly: (d) the **f bias world** model, which uses all observations, (e) the **p bias world** model, which weighs the observations by 1/2 and (f) the **n bias world** model, which uses the first half of the observations. In all cases, we used percentile calculations based on multiple runs and/or bootstrap replicates [11] to assess the significance of results.

3.1 Synthetic experiments

For this evaluation, we ran 10^4 independent experiments and employed multinomial models. For each experiment, we first generated the true prior distribution over user models γ. This was created by drawing Dirichlet parameters ϕ_i independently from a Gamma distribution. We also generated the true prior distribution over adversary models γ', by drawing from the same Gamma distribution. Then, a user model q was drawn from γ and an adversary model w was drawn from γ'. Finally, by flipping a coin, we generated data $x_1, \ldots, x_n$ from either q or w. Assuming equal prior probabilities of user and adversary, we predicted the most probable class and recorded the error. This was done for all subsequences of the observations' sequence x. Thus, the experiment measures the performance of methods when the amount of data that informs our decision increases.

For these experiments, we estimate the actual Dirichlet distribution with $\hat{\gamma}$. This estimation is performed via empirical Bayes using data from 1000 users drawn from the actual prior γ. At the k-th run, we draw a user model $q_k \sim \gamma$ and subsequently draw $x_k \sim q_k$. We then use $\hat{\gamma}$ and the user data $x_k \in \mathbb{N}^K$, to estimate a posterior over user models for the k-th user, $\psi_k(q) \triangleq \hat{\gamma}(q|x_k)$. The estimated prior $\hat{\gamma}$ is also used as the world model and as the prior over adversary models. The results, shown in Figure 1, show that the biased models consistently outperform the classic world model approach, while the partially biased models become significantly better than the fully biased models when the amount of observations increases. This is encouraging for application to real-world data.

3.2 Real data

The real world data were collected from an RFID based access control system used in two buildings of the TNO organization (Netherlands Organization for Applied Scientific Research). The data were collected during a three and a half month period, and they include successful accesses of 882 users, collected from 55 RFID readers granting access to users attempting to pass through doors in the buildings.

The initial data included three fields: the time and date that the access has been granted, the reader that has been used to get access and the ID of the RFID tag used[4]. In order to use the data in the experimental evaluation of the proposed model framework, we have discretized the time into hour-long intervals, and counted the number of accesses, per hour, per door for each user, in each day. This resulted in a total of $\approx 2 \cdot 10^5$ records. Since there are 24 hour-long slots in a day, and a total of 55 reader-equipped doors, this discretisation allowed us to model each user by a 1320-degree multinomial/Dirichlet model. Thus, even though the underlying Dirchlet/multinomial model framework is simple, the very high dimensionality of the observations makes the estimation and decision problem particularly taxing.

3.2.1 Experiments

We performed 10 independent runs. For the k-th run, we selected a random subset U_γ of the complete set of users U, such that $|U_\gamma|/|U| = 2/3$. We used U_γ to estimate the world model $\hat{\gamma}$. The remaining users $U_T = U \backslash U_\gamma$ were used to estimated the error rate over 10^3 repetitions. For the j-th repetition, we randomly selected a user $i \in U_T$ with at least 10 records D_i. We used half of those records, $\bar{D}_i$, to obtain $\psi_i(q) \triangleq \hat{\gamma}(q|\bar{D}_i)$. By flipping a coin, we obtain either (a) one record from $D_i \backslash \bar{D}_i$, or (b) data from some other user in U_T. Let us call that data x_j. For the biased models, we set $\xi = \hat{\gamma}$ and then used x_j to obtain $\xi(w|f(x_j))$, where $f(\cdot)$ denotes the appropriate transformation. Figure 2 shows results for the baseline world model approach (**world**), where $f(x) = \emptyset$, as the unmodified world model is used for the adversary, the full bias approach (**f bias**), where $f(x) = x$ since all the data is used, and finally the partial bias approach (**p bias**) where $f(x) = x/2$. The other approaches are not examined, as the oracle is not realisable, while the half-data and the all-but-last-data biased models are equivalent to the baseline world model, since we do not have a sequence of observations, but only a single record.

As can be seen in Figure 2, the baseline world model is always performing worse than the biased models, though in two runs the full bias model is close. Finally, though the two biased models are not distinguishable performance-wise, we noted a difference in the ratio of false positives to false negatives. Over the 10 runs, this was 0.2 ± 0.1 for the world model approach, 2.5 ± 0.5 for the fully biased model, and 0.9 ± 0.2 for the partially biased model.

4 Conclusion

We have presented a very simple, yet effective approach for classification problems where one class has no data. In particular, we define a prior over models which can be estimated from population data. This is adapted, as in the standard world-model approach, to a specific user. We introduce the idea of creating

[4]The data were sanitised to avoid privacy issues.

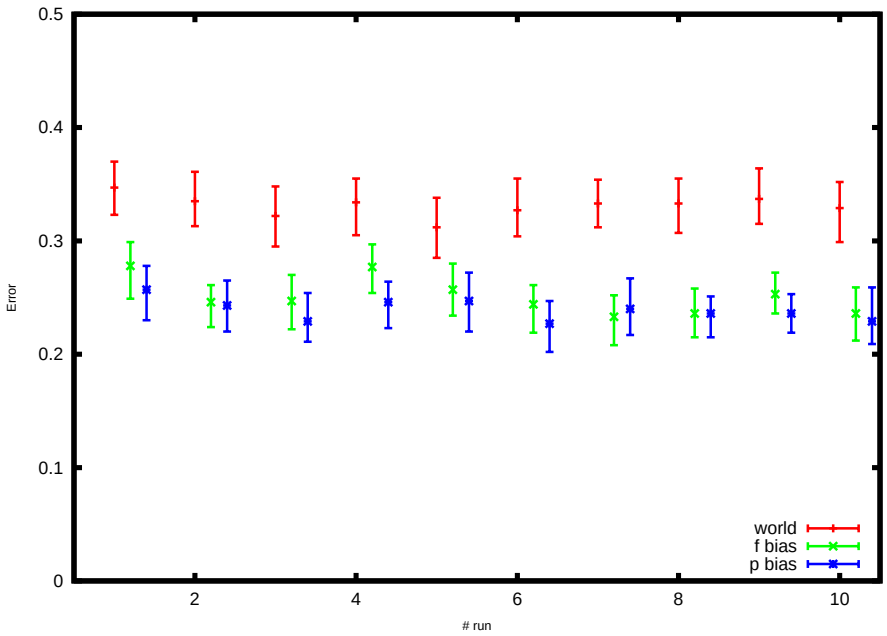

Figure 2: Error rates for 10 runs on the TNO door data. The error bars indicate top and bottom 5% percentiles from 100 bootstrap samples from 10^3 repetitions per run.

an adversary model, for which no *labelled* data exists, from the prior and currently seen data. Within the subjective Bayesian framework, this allows us to obtain a subjective upper bound on the probability of an attack.

Experimentally, it is shown [5] that: (a) we outperform the classical world model approach, while (b) it is always better to only *partially* condition the models on the new observations.

It is possible to extend the approach to the cost-sensitive case. Since we already have bounds on the probability of each class, together with a given cost matrix, we can also calculate bounds on the expected cost. This will allow us to make cost-sensitive decisions.

A related issue is whether to alter the *a priori class probabilities*; in our comparative experiments we used equal fixed values of 0.5. It is possible to utilise the population data to tune it in order to achieve some desired false positive / negative ratio. Such an automatic procedure would be useful for an expected performance curve [1] comparison between the various approaches. Finally, since the experiments on this relatively complex problem gave promising

[5] In an accompanying technical report [9], the effect of dimensionality on the performance of the method is also examined. There, it is shown that a Bayesian framework is essential for such a scheme to work and that naive approaches perform progressively worse as the dimensionality increases.

results, we plan to evaluate it on other problems that exhibit a lack of adversarial data.

A Proof

Lemma 2.1. For discrete $\mathcal{M}$, the marginal prior $\xi(x)$ can be re-written as follows:

$$\xi(x) = \sum_{\mu} \xi(x,\mu) = \sum_{\mu} \xi(x|\mu)\xi(\mu) = \sum_{\mu} \mu(x)\xi(\mu), \tag{5}$$

and similarly: $\xi'(x) = \frac{1}{\sum_\mu \mu(x)\xi(\mu)} \sum_\mu \mu(x)^2\xi(\mu)$. Thus, to prove the required statement, it is sufficient to show

$$\left(\sum_{\mu} \mu(x)^2\xi(\mu)\right)^{1/2} \geq \sum_{\mu} \mu(x)\xi(\mu). \tag{6}$$

Similarly, for continuous $\mathcal{M}$, we obtain:

$$\left(\int \mu(x)^2\, d\xi(\mu)\right)^{1/2} \geq \int \mu(x)\, d\xi(\mu). \tag{7}$$

In both cases, the norm induced by the probability measure ξ on $\mathcal{M}$ is $\|f\|_2 = (\int_{\mathcal{M}} |f(\mu)|^2\, d\xi(\mu))^{1/2}$, thus allowing us to included apply the Cauchy-Schwarz inequality $\|fg\|_1 \leq \|f\|_2\|g\|_2$. By setting $f(\mu) = \mu(x)$ and $g(\mu) = 1$, we obtain the required result, since $\|g\|_2 = (\int_{\mathcal{M}} d\xi(\mu))^{1/2} = 1$, as ξ is a probability measure. □

Acknowledgements

This work was supported by the Netherlands Organization for Scientific Research (NWO) under the RUBICON grant "Intrusion Detection in Ubiquitous Computing Technologies" and the ICIS project, supported by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024.

References

[1] S. Bengio, J. Mariéthoz, and M. Keller. The expected performance curve. In *ICML Workshop on ROC Analysis in MAchine Learning*, 2005.

[2] James O. Berger and Thomas Sellke. Testing a point null hypothesis: The irreconcilability of p values and evidence. *Journal of the American Statistical Association*, 82(397):112–122, Mar 1987.

[3] M. Breunig, H.-P. Kriegel, R.T. Ng, and J. Sander. LOF: Identifying density-based local outliers. *ACM SIGMOD Record*, 29(2):93–104, 2000.

[4] A.A. Cárdenas and J. D. Tygar. Statistical classification and computer security. In *Proceedings of the Workshop on Machine Learning in Adversarial Environments for Computer Security, (NIPS 2007)*, Whistler, BC, Canada, December 2007.

[5] F. Cardinaux, C. Sanderson, and S. Bengio. User authentication via adapted statistical models on face images. *IEEE Transactions on Signal Processing*, 54(1), January 2005.

[6] N. Dalvi, P. Domingos, Mausam, S. Sanghai, and D. Verma. Adversarial classification. In *10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 99–108, Seattle, WA, USA, 2004. ACM Press.

[7] M. H. DeGroot. *Optimal Statistical Decisions*. John Wiley & Sons, 1970, Republished in 2004.

[8] H. Deng, Q. Zeng, and D.P. Agrawal. SVM-based intrusion detection system for wireless ad hoc networks. In *Proceedings of the 58th IEEE Vehicular Technology Conference (VTC'03)*, volume 3, pages 2147–2151, Orlando, FL, USA, 6-9 October 2003.

[9] Christos Dimitrakakis and Aikaterini Mitrokotsa. Statistical decision making for authentication and intrusion detection. Technical Report IAS-UVA-09-02, University of Amsterdam, April 2009.

[10] W. Edwards, H. Lindman, and L. J. Savage. Bayesian statistical inference for psychological research. *Psychological Review*, 70:193–242, 1963.

[11] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*, volume 57 of *Monographs on Statistics & Applied Probability*. Chapmann & Hall, November 1993. ISBN 0412042312.

[12] S. Furui. Robust speech recognition. In K. Ponting, editor, *Computational Models of Speech Pattern Processing*, NATO ASI Series, pages 132–142. Springer-Verlag, Berlin, 1999.

[13] J. Gauvain and Chin-Hui Lee. Maximum a posteriori estimation for multivariate gaussian mixture observations of markov chains. *IEEE Transactions on Speech and Audio Processing*, 2:291–298, 1994.

[14] Y. Liu, Y. Li, and H. Man. MAC layer anomaly detection in ad hoc networks. In *Proceedings of the 6th Annual IEEE SMC Information Assurance Workshop (IAW '05)*, pages 402–409, West Point, NY, USA, 15-17 June 2005.

[15] D. Lowd and C. Meek. Adversarial learning. In *Proceedings of the 11th ACM International on Knowledge Discovery and Data Mining (ACM SIGKDD '05)*, pages 641–647, Chicago, IL, USA, 2005.

[16] T. Minka. Estimating a Dirichlet distribution, 2003.

[17] S. Mukkamala, A.H. Sung, and B. Abraham. Intrusion detection using an ensemble of intelligent paradigms. *Journal of Network and Computer Applications, Special Issue on Computational Intelligence on the Internet*, 28(2):167–182, 2005.

[18] L. Portnoy, E. Eskin, and S.J. Stolfo. Intrusion detection with unlabeled data using clustering. In *Proceedings of the ACM CSS Workshop on Data Mining Applied to Security (DMSA '01)*, pages 5–8, Philadelphia, PA, USA, 5-8 November 2001.

[19] S. Ramaswamy, R. Rastogi, and K. Shim. Efficient algorithms for mining outliers from large data sets. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 427–438, Dallas, TX, USA, 14-19 May 2000.

[20] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn. Speaker verification using adapted gaussian mixture models. *Digital Signal Processing*, 10:19–41, 2000.

[21] H. Robbins. An empirical bayes approach to statistics. In J. Neyman, editor, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, pages 157–163. University of California Press, Berkeley, CA, 1955.